

Long-time stability and accuracy of the ensemble Kalman-Bucy filter for fully observed processes and small measurement noise

Jana de Wiljes*

Sebastian Reich†

Wilhelm Stannat‡

March 1, 2017

Abstract

The ensemble Kalman filter has become a popular data assimilation technique in the geosciences. However, little is known theoretically about its long term stability and accuracy. In this paper, we investigate the behavior of an ensemble Kalman-Bucy filter applied to continuous-time filtering problems. We derive mean field limiting equations as the ensemble size goes to infinity as well as uniform-in-time accuracy and stability results for finite ensemble sizes. The later results require that the process is fully observed and that the measurement noise is small. We also demonstrate that our ensemble Kalman-Bucy filter is consistent with the classic Kalman-Bucy filter for linear systems and Gaussian processes. We finally verify our theoretical findings for the Lorenz-63 system.

Keywords. Data assimilation, Kalman-Bucy filter, ensemble Kalman filter, stability, accuracy, asymptotic behavior

AMS(MOS) subject classifications. 65C05, 62M20, 93E11, 62F15, 86A22

1 Introduction

In this paper, we consider the continuous-time filtering problem [Jaz70, BC08] for diffusion processes of type

$$dX_t = f(X_t)dt + \sqrt{2}C dW_t \quad (1)$$

and observations, Y_t , given by

$$dY_t = h(X_t)dt + R^{1/2}dV_t. \quad (2)$$

Here X_t denotes the state variable of the N_x -dimensional diffusion process with Lipschitz-continuous drift $f : \mathbb{R}^{N_x} \rightarrow \mathbb{R}^{N_x}$ and constant diffusion tensor $D = CC^T$ and $C \in \mathbb{R}^{N_x \times N_w}$. The observations Y_t are N_y -dimensional with forward map $h : \mathbb{R}^{N_x} \rightarrow \mathbb{R}^{N_y}$ and measurement error covariance matrix $R \in \mathbb{R}^{N_y \times N_y}$. Finally, $W_t \in \mathbb{R}^{N_w}$ and $V_t \in \mathbb{R}^{N_y}$ denote independent Brownian motion of dimension N_w and N_y , respectively. It is well-known that the filtering distribution π_t , i.e., the conditional distribution in X_t for given observations Y_s , $s \in [0, t]$, satisfies the Kushner-Zakai equation [Jaz70, BC08], which we write as an evolution equation in the expectation values

$$\pi_t[g] = \int_{\mathbb{R}^{N_x}} g(x) \pi_t(x) dx \quad (3)$$

of smooth and bounded functions $g : \mathbb{R}^{N_x} \rightarrow \mathbb{R}$, i.e.

$$d\pi_t[g] = \pi_t[f \cdot \nabla g]dt + \pi_t[\nabla \cdot D \nabla g]dt + (\pi_t[gh] - \pi_t[g]\pi_t[h])^T R^{-1} (dY_t - \pi_t[h]dt). \quad (4)$$

In order to have a properly formulated filtering problem, we also have to specify the distribution at initial time $t = 0$.

*Universität Potsdam, Institut für Mathematik, Karl-Liebknecht-Str. 24/25, D-14476 Potsdam, Germany

†Universität Potsdam, Institut für Mathematik, Karl-Liebknecht-Str. 24/25, D-14476 Potsdam, Germany (sreich@math.uni-potsdam.de) and University of Reading, Department of Mathematics and Statistics, Whiteknights, PO Box 220, Reading RG6 6AX, UK

‡TU Berlin, Institut für Mathematik, Str. des 17. Juni 136, D-10623 Berlin, Germany

Popular numerical methods for approximating solutions to (4) include direct finite-difference or finite-element discretizations of (4) and sequential Monte Carlo methods, also called particle filters [BC08, DdFe01]. These methods lead to consistent approximations but are typically restricted to low-dimensional problems. In recent years, particle filter methods have become popular, which are applicable to higher-dimensional problems but are no longer consistent. These include the ensemble Kalman filter (EnKF) [Eve06, LSZ15, RC15], which is now widely used in the geosciences.

Abstractly spoken, particle filters are defined as follows. First one defines M weighted random variables X_t^i , called particles, which are i.i.d. at initial time $t = 0$ with distribution π_0 , and weights $w_t^i \geq 0$ with $w_0^i = 1/M$ at initial time. A particle filter is then characterized by appropriate evolution laws for the particles and the weights. Most known particle filters lead to particles which remain identically distributed while no longer being independent, so called interacting particle systems [Mor13]. If the weights are furthermore kept uniform either through resampling or appropriately defined evolution equations, then expectation can be taken with respect to the law π_t^M of the M th particle and consistency of a particle filter means that $\lim_{M \rightarrow \infty} \pi_t^M[g] \rightarrow \pi_t[g]$.

The classic bootstrap filter [AMGC02] uses (1) for the evolution of the particles and (2) for the update of the weights in combination with an appropriate resampling strategy in order to avoid the weights to degenerate. The EnKF, on the contrary, introduces modified evolution equations for the particles which include the observations and keep the weights uniform instead. Most available EnKF formulations are stated for discrete-in-time observations [Eve06]. While the robust behavior of EnKFs has been demonstrated for many applications primarily arising from the geosciences, our theoretical understanding of their long-time stability and accuracy is still rather limited. Large sample size limits have been, for example, investigated in [GMT11, KM15] and it has been demonstrated that the EnKF converges to the classic Kalman filter for linear systems (1), linear observations (2) and Gaussian initial conditions. Using concepts from shadowing, [GTH13] showed that the EnKF is stable and accurate uniformly in time for hyperbolic dynamical systems provided the ensemble size is larger than the dimension of the chaotic attractor. Stability and ergodicity of EnKFs have also been studied in [TMK16]. The authors demonstrate that the extended system consisting of (1), (2), and the filter algorithm possesses a unique ergodic invariant measure provided the existence of an appropriate Lyapunov function can be guaranteed. While such ergodicity results of [MH12] are important, they do not imply accuracy of a filter. In fact, it is well known, that ensemble Kalman filter can diverge and techniques, such as ensemble inflation [Eve06], have been developed in order to stabilize a filter. In fact, it has been rigorously demonstrated, for example, in [KLS14] that a judicious choice of inflation can lead to uniform-in-time accurate state estimates.

In this paper, we investigate a time-continuous EnKF formulation which is consistent with the classic Kalman filter in the linear case and which is also stable and accurate uniformly in time without additional ensemble inflation. In this first study, we will assume for simplicity that the system is fully observable and the measurement errors are small. These assumptions can be relaxed under appropriate assumptions on the stochastic process (1) and the observation process (2), well known from the theory of classic Kalman filter theory (i.e. observability and controllability) [Jaz70]. These generalization will be studied in a future publication.

The specific ensemble Kalman-Bucy filter (EnKBF) formulation, which we will investigate in this paper, is given by drawing M independent realizations (called particles or ensemble members) $X_0^i \sim \pi_0$, which then follow the system of differential equations

$$dX_t^i = f(X_t^i)dt + D(P_t^M)^{-1}(X_t^i - \bar{x}_t^M)dt - \frac{1}{2}Q_t^M R^{-1}(h(X_t^i)dt + \bar{h}_t^M dt - 2dY_t) \quad (5)$$

for $t \geq 0$. These equations of motion for the particles are closed through the empirical estimates

$$\bar{x}_t^M = \frac{1}{M} \sum_{i=1}^M X_t^i, \quad P_t^M = \frac{1}{M-1} \sum_{i=1}^M (X_t^i - \bar{x}_t^M)(X_t^i - \bar{x}_t^M)^T, \quad (6)$$

and

$$\bar{h}_t^M = \frac{1}{M} \sum_{i=1}^M h(X_t^i), \quad Q_t^M = \frac{1}{M-1} \sum_{i=1}^M (X_t^i - \bar{x}_t^M)(h(X_t^i) - \bar{h}_t^M)^T. \quad (7)$$

In case P_t^M is not invertible, which is surely the case for $M \leq N_x$, the inverse of P_t^M is replaced by its generalized inverse $(P_t^M)^+$. This generalization is unproblematic since $(P_t^M)^+$ gets multiplied by a vector which is in the range of P_t^M . Finally, given a solution of (5), we define the empirical expectation values of a function g and the empirical distribution $\hat{\pi}_t^M$ by

$$\bar{g}_t^M := \hat{\pi}_t^M[g], \quad \hat{\pi}_t^M(x) := \frac{1}{M} \sum_{i=1}^M \delta(x - X_t^i), \quad (8)$$

respectively. Here $\delta(\cdot)$ denotes the standard Dirac delta function. The formulation (5) has been stated first in [BR10, BR12]. Alternative ensemble Kalman-Bucy formulations include stochastically perturbed formulations [Rei11, LSZ15, RC15] and the extended ensemble Kalman-Bucy filter, whose exponential stability and propagation of chaos properties have been studied in [DMKT16].

Given the evolution equations (5), one can derive associated evolution equations for the ensemble mean, \bar{x}_t^M , and the ensemble covariance matrix, P_t^M . These are given by

$$d\bar{x}_t^M = \bar{f}_t^M dt - Q_t^M R^{-1}(\bar{h}_t^M dt - dY_t) \quad (9)$$

with $\bar{f}_t^M = \hat{\pi}_t^M[f]$ and

$$\begin{aligned} \frac{d}{dt} P_t^M &= \frac{1}{M-1} \sum_{i=1}^M \{ (f(X_t^i) - \bar{f}_t)(X_t^i - \bar{x}_t)^T + (X_t^i - \bar{x}_t)(f(X_t^i) - \bar{f}_t)^T \} + \\ &+ \{ D(P_t^M)^+ P_t^M + P_t^M (P_t^M)^+ D \} - Q_t^M R^{-1} (Q_t^M)^T. \end{aligned} \quad (10)$$

We will study the behavior of the EnKBF for fully observed processes, i.e. $h(x) = x$ and small measurement noise, i.e. $R = \varepsilon I$, in Sections 2 and 3. More specifically, it is shown in Section 2 that strong solutions of (5) exist for all times and are unique. Uniform-in-time stability and accuracy of (5) are proven in Section 3. In Sections 4 and 5, we return to the filtering problem for general observation operator, h , and measurement error covariance matrix R . It is demonstrated in Section 4 that (9) and (10) are consistent with the classic Kalman-Bucy for linear systems [Jaz70]. We study the asymptotic mean field limiting equations for (5) as $M \rightarrow \infty$ in Section 5. More specifically, these mean field limiting equations are given by

$$d\hat{X}_t = f(\hat{X}_t)dt + D(\mathcal{P}_t)^{-1}(\hat{X}_t - \bar{x}_t)dt - \frac{1}{2} \mathcal{Q}_t R^{-1} \left(h(\hat{X}_t)dt + \bar{h}_t dt - 2dY_t \right), \quad (11)$$

with $\bar{x}_t = \hat{\pi}_t[x]$, $\bar{h}_t = \hat{\pi}_t[h]$,

$$\mathcal{P}_t = \text{Cov}(\hat{X}_t, \hat{X}_t), \quad \mathcal{Q}_t = \text{Cov}(\hat{X}_t, h(\hat{X}_t)). \quad (12)$$

Here $\hat{\pi}_t$ denotes the distribution of \hat{X}_t . Using Itô's formula, it follows that the evolution equation for the distribution $\hat{\pi}_t$ is different from the Kushner-Zakai equation (4) and is instead given by

$$d\hat{\pi}_t[g] = \hat{\pi}_t \left[\nabla g \cdot \left\{ f dt + D\mathcal{P}_t^{-1}(x - \hat{\pi}_t[x])dt - \frac{1}{2} \mathcal{Q}_t R^{-1}(h(x)dt + \hat{\pi}_t[h]dt - 2dY_t) \right\} + \frac{1}{2} \nabla \cdot \mathcal{Q}_t R^{-1} \mathcal{Q}_t^T \nabla g dt \right]. \quad (13)$$

Some numerical results, supporting our theoretical estimates, will be presented in Section 6 using a stochastically perturbed Lorenz-63 system [Lor63, LSZ15].

2 Well-posedness of the ensemble Kalman-Bucy filter for fully observed processes

In this section, we specify the problem setting which is investigated in detail in this paper. We will also derive a first well-posedness result for the system (5) -(7). More specifically, we assume that the process is fully observed, i.e. $h(x) = x$, that the covariance matrix of the measurement noise is of the form $R = \varepsilon I$ with $\varepsilon > 0$ and that the diffusion tensor D has full rank. Since the ensemble size, M , will be fixed in this section, we also drop the superscript M in (5). Hence (5) is replaced by

$$dX_t^i = f(X_t^i)dt + DP_t^+(X_t^i - \bar{x}_t)dt - \frac{1}{2\varepsilon} P_t (X_t^i dt + \bar{x}_t dt - 2dY_t), \quad (14)$$

$i = 1, \dots, M$. Inner product in R^{N_x} will be denoted by $\langle \cdot, \cdot \rangle$ and we recall that

$$\langle a, b \rangle = \text{tr}(ba^T). \quad (15)$$

Hence we quickly verify that

$$\frac{1}{M-1} \sum_{i=1}^M \langle X_t^i - \bar{x}_t, DP_t^+(X_t^i - \bar{x}_t) \rangle = \text{tr}(DP_t^+ P_t) \quad (16)$$

and

$$\frac{1}{M-1} \sum_{i=1}^M \langle X_t^i - \bar{x}_t, P_t(X_t^i - \bar{x}_t) \rangle = \text{tr}(P_t^2) = \|P_t\|_F^2. \quad (17)$$

Here $\|A\|_F$ denotes the Frobenius norm of a matrix A . We also introduce the notation $\langle A, B \rangle = \text{tr}(BA^T)$, i.e. $\|A\|_F^2 = \langle A, A \rangle$.

We now investigate the l_2 -norm of the ensemble deviations from the mean, i.e.

$$V_t = \frac{1}{M-1} \sum_{i=1}^M \|X_t^i - \bar{x}_t\|^2, \quad (18)$$

which satisfies the evolution equation

$$\begin{aligned} \frac{1}{2} \frac{dV_t}{dt} &= \frac{1}{M-1} \sum_{i=1}^M \langle X_t^i - \bar{x}_t, f(X_t^i) - \bar{f}_t \rangle + \frac{1}{M-1} \sum_{i=1}^M \langle X_t^i - \bar{x}_t, DP_t^+(X_t^i - \bar{x}_t) \rangle - \\ &\quad - \frac{1}{2\varepsilon} \frac{1}{M-1} \sum_{i=1}^M \langle X_t^i - \bar{x}_t, P_t(X_t^i - \bar{x}_t) \rangle \\ &= \frac{1}{M-1} \sum_{i=1}^M \langle X_t^i - \bar{x}_t, f(X_t^i) - \bar{f}_t \rangle + \text{tr}(DP_t^+ P_t) - \frac{1}{2\varepsilon} \text{tr}(P_t^2) \\ &= \frac{1}{M-1} \sum_{i=1}^M \langle X_t^i - \bar{x}_t, f(X_t^i) - f(\bar{x}_t) \rangle + \text{tr}(DP_t^+ P_t) - \frac{1}{2\varepsilon} \|P_t\|_F^2 \end{aligned} \quad (19)$$

Here we have used

$$\sum_i \langle X_t^i - \bar{x}_t, f(\bar{x}_t) - f_t \rangle = 0 \quad (20)$$

and that the evolution equation (9) for the mean, \bar{x}_t , reduces to

$$d\bar{x}_t = \bar{f}_t dt - \frac{1}{\varepsilon} P_t(\bar{x}_t dt - dY_t) \quad (21)$$

in our setting.

Lemma 2.1. *The Frobenius norm of P_t satisfies*

$$\frac{1}{\sqrt{M}} V_t \leq \|P_t\|_F \leq V_t. \quad (22)$$

Proof. We first note the following identity:

$$\|P_t\|_F^2 = \frac{1}{(M-1)^2} \sum_{i,j} \langle X_t^i - \bar{x}_t, X_t^j - \bar{x}_t \rangle^2. \quad (23)$$

For the proof of the upper bound it is now sufficient to observe that

$$\begin{aligned} \frac{1}{(M-1)^2} \sum_{i,j} \langle X_t^i - m_t, X_t^j - \bar{x}_t \rangle^2 &\leq \frac{1}{(M-1)^2} \sum_{i,j} \|X_t^i - \bar{x}_t\|^2 \|X_t^j - \bar{x}_t\|^2 \\ &= \left(\frac{1}{M-1} \sum_i \|X_t^i - \bar{x}_t\|^2 \right)^2. \end{aligned} \quad (24)$$

For the proof of the lower bound observe that

$$\frac{1}{(M-1)^2} \sum_{i,j} \langle X_t^i - m_t, X_t^j - \bar{x}_t \rangle^2 \geq \frac{1}{(M-1)^2} \sum_i \|X_t^i - \bar{x}_t\|^4 \geq \frac{1}{M} \left(\frac{1}{M-1} \sum_i \|X_t^i - \bar{x}_t\|^2 \right)^2. \quad (25)$$

□

Remark 2.2. We recall the standard relations between the Frobenius and the spectral norm of a matrix, i.e.,

$$\|P_t\| \leq \|P_t\|_F \quad (26)$$

and

$$\|P_t\|_F \leq \sqrt{N_x} \|P_t\|. \quad (27)$$

We are now ready to obtain uniform-in-time upper and lower bounds on V_t . First, we can estimate the first term of (19) from above and from below as follows:

$$\frac{1}{M-1} \sum_i \langle X_t^i - \bar{x}_t, f(X_t^i) - f(\bar{x}_t) \rangle \leq L_+ \frac{1}{M-1} \sum_i \|X_t^i - \bar{x}_t\|^2 = L_+ V_t \quad (28)$$

and

$$\frac{1}{M-1} \sum_i \langle X_t^i - \bar{x}_t, f(X_t^i) - f(\bar{x}_t) \rangle \geq L_- \frac{1}{M-1} \sum_i \|X_t^i - \bar{x}_t\|^2 = L_- V_t, \quad (29)$$

respectively, where

$$L_+ := \sup_{x \neq y} \frac{\langle f(x) - f(y), x - y \rangle}{\|x - y\|^2} \quad L_- := \inf_{x \neq y} \frac{\langle f(x) - f(y), x - y \rangle}{\|x - y\|^2} \quad (30)$$

are the upper and lower control on the “dissipativity” constant of f . For Lipschitz continuous f with Lipschitz constant L we clearly have $L_+ \leq L$ and $L_- \geq -L$. Provided $V_t \neq 0$, we also find that

$$d_0 \leq \text{tr}(DP_t^+ P_t) \leq \text{tr}(D), \quad (31)$$

where $d_0 \geq 0$ denotes the smallest diagonal element of D .

Inserting these estimates and the previous two identities into (19) we first obtain the upper bound

$$\frac{1}{2} \frac{dV_t}{dt} \leq L_+ V_t + \text{tr}(D) - \frac{1}{2\varepsilon M} V_t^2. \quad (32)$$

This implies that $V_t \leq \max\{V_0, \varepsilon L_+ M + \sqrt{(\varepsilon M L_+)^2 + 2\varepsilon M \text{tr}(D)}\}$ uniformly in t . Similarly, we obtain the lower bound

$$\frac{1}{2} \frac{dV_t}{dt} \geq L_- V_t + d_0 - \frac{1}{2\varepsilon} V_t^2, \quad (33)$$

which implies that $V_t \geq \min\{V_0, \varepsilon L_- + \sqrt{(\varepsilon L_-)^2 + 2\varepsilon d_0}\}$ uniformly in t and $V_t > 0$ provided $V_0 > 0$.

Theorem 2.3. Assume that the drift term f in (1) is globally Lipschitz continuous and satisfies a linear growth condition

$$\|f(x)\| \leq \tilde{c}_1(1 + \|x\|) \quad (34)$$

for an appropriate $\tilde{c}_1 > 0$. Then the system (14) together with (6)-(7) possesses strong solutions for all times $t \geq 0$.

Proof. We can decompose the equations (14) into ordinary differential equations in $X_t^i - \bar{x}_t$, $i = 1, \dots, M$ and Equation (21) for the mean, \bar{x}_t . Since the l_2 -norm, V_t , remains bounded, the equations in $X_t^i - \bar{x}_t$ are well-posed. Furthermore, since $\|P_t\|$ remains bounded as well, the combined drift term in (21), written as

$$d\bar{x}_t = f(\bar{x}_t)dt + b_t dt - \frac{1}{\varepsilon} P_t(\bar{x}_t dt - dY_t), \quad (35)$$

with $b_t = \bar{f}_t - f(\bar{x}_t)$, is Lipschitz continuous in \bar{x}_t and, hence, satisfies a linear growth condition, i.e.

$$\|f(\bar{x}_t) + b_t - \varepsilon^{-1} P_t \bar{x}_t\| \leq \|f(\bar{x}_t) - \varepsilon^{-1} P_t \bar{x}_t\| + \|\bar{f}_t - f(\bar{x}_t)\| \leq \tilde{c}_2(1 + \|\bar{x}_t\|) \quad (36)$$

for an appropriate $\tilde{c}_2 > 0$, and, hence, strong solutions to (21) exist for all times [Øks00]. \square

Remark 2.4. For the analysis of the asymptotic behavior of $M \rightarrow \infty$ the upper bound on the V_t is not sufficient, because it diverges as $M \rightarrow \infty$. However, since we need a control only locally in time, we can use (32) to estimate $\frac{1}{2} \frac{d}{dt} V_t \leq L_+ V_t + \text{tr}(D)$ which implies the upper bound

$$V_t \leq e^{2L_+ t} \left(V_0 + \frac{\text{tr}(D)}{L_+} \right), \quad (37)$$

which becomes uniform in M (but of course not in t) if the particles at time $t = 0$ are chosen with uniform upper bound on $V_0 = V_0^M$.

3 Accuracy of the ensemble Kalman-Bucy filter for finite ensemble sizes and small measurement noise

The goal of this section is to derive bounds on the estimation error

$$e_t = X_t^{\text{ref}} - \bar{x}_t, \quad (38)$$

where X_t^{ref} denotes the reference trajectory of (1) which generated the data. We again restrict the discussion to fully observed processes and small measurement errors, i.e.,

$$dY_t = X_t^{\text{ref}} dt + \sqrt{\varepsilon} dV_t \quad (39)$$

and drop the superscript M from all relevant quantities, as we are interested in the accuracy of the filter for fixed ensemble size, M .

We find that the estimation error satisfies the evolution equation

$$de_t = f(X_t^{\text{ref}})dt + \sqrt{2}C dW_t - \bar{f}_t dt - \frac{1}{\varepsilon} P_t (e_t dt - \varepsilon^{1/2} dV_t). \quad (40)$$

Then Ito's formula implies that

$$\begin{aligned} \frac{1}{2} d\|e_t\|^2 &= \langle f(X_t^{\text{ref}}) - \bar{f}_t, X_t^{\text{ref}} - \bar{x}_t \rangle dt - \frac{1}{\varepsilon} \langle e_t, P_t e_t \rangle dt \\ &\quad + \langle e_t, \sqrt{2} dW_t \rangle + \frac{1}{\sqrt{\varepsilon}} \langle e_t, P_t dV_t \rangle + \text{tr}(D) dt + \frac{1}{2\varepsilon} \text{tr}(P_t^2) dt. \end{aligned}$$

We introduce $E_t = \|e_t\|^2/2$ and write

$$dE_t = \mathcal{E}_t dt + dM_t \quad (41)$$

with

$$\mathcal{E}_t = \langle X_t^{\text{ref}} - \bar{x}_t, f(X_t^{\text{ref}}) - \bar{f}_t \rangle - \frac{1}{\varepsilon} \langle e_t, P_t e_t \rangle + \text{tr}(D) + \frac{1}{2\varepsilon} \|P_t\|_F^2 \quad (42)$$

and the martingale

$$M_t = \int_0^t \langle e_s, \varepsilon^{-1/2} P_s dV_s + \sqrt{2} dW_s \rangle, t \geq 0.$$

To make further progress we need bounds for the smallest and largest singular values of P_t denoted by λ_t^{\min} and λ_t^{\max} , respectively. An upper bound for the largest singular value has already been derived in Section 2, since $\lambda_t^{\max} = \|P_t\| \leq \|P_t\|_F \leq V_t$. Let us assume for now that P_t is invertible, then the explicit evolution equation for P_t reduces to

$$\frac{d}{dt} P_t = \frac{1}{M-1} \sum_i (f(X_t^i) - \bar{f}(t))(X_t^i - \bar{x}_t)^T + (X_t^i - \bar{x}_t)(f(X_t^i) - \bar{f}(t))^T + 2D - \frac{1}{\varepsilon} P_t^2. \quad (43)$$

Next we make use of the fact that P_t can be diagonalized, i.e., there are orthogonal matrices Q_t and diagonal matrices Λ_t such that

$$P_t = Q_t^T \Lambda_t Q_t. \quad (44)$$

While the orthogonal matrices Q_t are in general only continuous in t , the diagonal matrix of singular values can be chosen to be differentiable in t [Rel69]. As shown in [DE99], the evolution equation for diagonal matrix of eigenvalues, Λ_t , is of the form

$$\frac{d\Lambda_t}{dt} = \text{diag}(Q_t U_t Q_t^T) + 2\text{diag}(Q_t D Q_t^T) - \frac{1}{\varepsilon} \Lambda_t^2 \quad (45)$$

with

$$U_t := \frac{1}{M-1} \sum_i \{f(X_t^i) - f(\bar{x}_t)\} \{X_t^i - \bar{x}_t\}^T + \{X_t^i - \bar{x}_t\} \{f(X_t^i) - f(\bar{x}_t)\}^T. \quad (46)$$

Here $\text{diag}(A)$ denotes a diagonal matrix with diagonal entries equal to the diagonal of A . More specifically, the diagonal entries of $\text{diag}(Q_t U_t Q_t^T)$ are given by

$$(\text{diag}(Q_t U_t Q_t^T))_{ii} = e_i^T Q_t U_t Q_t^T e_i \quad (47)$$

where $e_i \in \mathbb{R}^{N_x}$ denotes the i th basis vector in \mathbb{R}^{N_x} .

Next we derive the following estimate using the fact that f is Lipschitz continuous with Lipschitz constant $L > 0$. Then, given any unit vector v , it holds that

$$\left| \frac{1}{M-1} \sum_i \langle f(X_t^i) - f(\bar{x}_t), v \rangle \langle X_t^i - \bar{x}_t, v \rangle \right| \leq \left(\frac{1}{M-1} \sum_i \langle f(X_t^i) - f(\bar{x}_t), v \rangle^2 \right)^{\frac{1}{2}} \left(\frac{1}{M-1} \sum_i \langle X_t^i - \bar{x}_t, v \rangle^2 \right)^{\frac{1}{2}} \quad (48)$$

$$\leq L V_t \quad (49)$$

$$\leq L \sqrt{N_x M} \|P_t\|, \quad (50)$$

where we have used $V_t \leq \sqrt{N_x M} \|P_t\|$.

Hence setting $v = Q_t^T e_i$, we obtain

$$|(\text{diag}(Q_t U_t Q_t^T))_{ii}| \leq 2L \sqrt{N_x M} \|P_t\| = 2L \sqrt{N_x M} \lambda_t^{\max}. \quad (51)$$

Since $\lambda_t^{\max} = (\Lambda_t)_{ii}$ for some index i , we hence deduce that

$$\frac{d\lambda_t^{\max}}{dt} \leq 2L \sqrt{N_x M} \lambda_t^{\max} + 2d_{\max} - \frac{(\lambda_t^{\max})^2}{\varepsilon}, \quad (52)$$

where d_{\max} denotes the largest eigenvalue of D . This implies that

$$\lambda_t^{\max} \leq \max \left\{ \lambda_0^{\max}, \varepsilon L \sqrt{N_x M} + \sqrt{\varepsilon^2 L^2 N_x M + 2\varepsilon d_{\max}} \right\}. \quad (53)$$

Hence we have shown the following

Lemma 3.1. (upper bound on spectral radius of P_t) *There is a constant $C_1 = C_1(L, M, N_x, D, \varepsilon_0)$ such that $\lambda_0^{\max} \leq C_1 \varepsilon^{1/2}$ at initial time $t = 0$ implies $\lambda_t^{\max} \leq C_1 \varepsilon^{1/2}$ for all times and all $\varepsilon \leq \varepsilon_0$.*

We now use our upper bound on $\lambda_t^{\max} = \|P_t\|_2$ from Lemma 3.1 in order to get the estimate

$$|(\text{diag}(Q_t U_t Q_t^T))_{ii}| \leq 2L \sqrt{N_x M} C_1 \varepsilon^{1/2}. \quad (54)$$

Hence, we deduce that

$$\frac{d\lambda_t^{\min}}{dt} \geq -2L \sqrt{N_x M} C_1 \varepsilon^{1/2} + 2d_{\min} - \frac{(\lambda_t^{\min})^2}{\varepsilon} \quad (55)$$

and

$$\lambda_t^{\min} \geq \min \left\{ \lambda_0^{\min}, -\varepsilon^{3/2} L C_1 \sqrt{N_x M} + \sqrt{\varepsilon^3 L^2 C_1^2 N_x M + 2\varepsilon d_{\min}} \right\}, \quad (56)$$

which implies the desired lower bound on λ_t^{\min} . Here d_{\min} denotes the smallest eigenvalue of D . We now fix $\varepsilon_0 > 0$ such that

$$-\varepsilon_0^{3/2} L C_1 \sqrt{N_x M} + \sqrt{\varepsilon_0^3 L^2 C_1^2 N_x M + \varepsilon_0 d_{\min}} > 0. \quad (57)$$

Lemma 3.2. (lower bound on smallest singular value of P_t) *There is a constant $C_2 = C_2(L, M, N_x, D, \varepsilon_0)$ such that $\lambda_0^{\min} \geq C_2 \varepsilon^{1/2}$ at initial time $t = 0$ implies $\lambda_t^{\min} \geq C_2 \varepsilon^{1/2}$ for all $t > 0$ and all $\varepsilon \leq \varepsilon_0$.*

Remark 3.3. *The upper and lower bounds for the largest and smallest, respectively, eigenvalue of P_t depend on the ensemble size, M . This dependence can be eliminated for the price of the estimates no longer being valid uniformly in time. We now derive such M -independent upper and lower bounds. Let us assume that*

$$2LV_s^{1/2} (\lambda_s^{\max})^{1/2} \leq d_{\min} \quad (58)$$

for all $s \in [0, t]$. Such a bound can be found because of (37) and for ε sufficiently small, i.e. $\varepsilon \leq \varepsilon_t$. Then (52) implies that

$$\lambda_s^{\max} \leq 2d_{\max}^{1/2}\varepsilon^{1/2} \quad (59)$$

for all $s \in [0, t]$ and all $\varepsilon \leq \varepsilon_t$. Similarly, (55) implies that

$$\lambda_s^{\min} \geq d_{\min}^{1/2}\varepsilon^{1/2}. \quad (60)$$

Hence we have traded the M -dependent constants C_1 and C_2 in the previous two lemmas by M -independent constants $\tilde{C}_1 = 2d_{\max}^{1/2}$ and $\tilde{C}_2 = d_{\min}^{1/2}$, respectively. However, the estimates hold for $\varepsilon \leq \varepsilon_t$ only, where the upper bound $\varepsilon_t = \varepsilon_t(L, D)$ decreases in time.

The upper and lower bounds of the eigenvalues of P_t obtained in the previous two lemmas hold with constants C_1 and C_2 independent of the driving Wiener processes. They only depend on the initial conditions (which might be random), but we can impose deterministic bounds on the spectral radius of the covariance matrix. Hence we can take expectations on both sides of (41) in order to obtain the following integral inequality

$$\begin{aligned} \mathbb{E}[E_t] &\leq \mathbb{E}[E_0] + \int_0^t \mathbb{E}[\mathcal{E}_s] \, ds \\ &\leq \mathbb{E}[E_0] + \int_0^t \mathbb{E} \left[\text{tr}(D) + N_x C_1^2 + 2\varepsilon^{1/4} L C_1^{1/2} (N_x M)^{1/4} E_s^{1/2} - 2 \frac{C_2 - \varepsilon^{1/2} L}{\varepsilon^{1/2}} E_s \right] \, ds, \end{aligned} \quad (61)$$

where we used

$$\begin{aligned} \langle f(X_t^{\text{ref}}) - \bar{f}_t, X_t^{\text{ref}} - \bar{x}_t \rangle &= \langle f(X_t^{\text{ref}}) - f(\bar{x}_t), X_t^{\text{ref}} - \bar{x}_t \rangle + \langle f(\bar{x}_t) - \bar{f}_t, X_t^{\text{ref}} - \bar{x}_t \rangle \\ &\leq 2L_+ E_t + \sqrt{2} L V_t^{1/2} E_t^{1/2} \\ &\leq 2L \left(E_t + \varepsilon^{1/4} C_1^{1/2} (M N_x)^{1/4} E_t^{1/2} \right). \end{aligned} \quad (62)$$

The next step is to close the right hand side in $\mathbb{E}[E_s]$. To this end, we first derive the following ω -wise estimate

$$\begin{aligned} \mathcal{E}_s &\leq \left(\text{tr}(D) + N_x C_1^2 + 2\varepsilon^{1/4} L C_1^{1/2} (N_x M)^{1/4} E_s^{1/2} - 2 \frac{C_2 - \varepsilon^{1/2} L}{\varepsilon^{1/2}} E_s \right) \\ &\leq C_3 + \varepsilon^{1/4} C_4 E_s^{1/2} - 2 \frac{C_2 - \varepsilon^{1/2} L}{\varepsilon^{1/2}} E_s \\ &\leq \left(C_3 + \varepsilon \frac{C_4^2}{C_2} \right) - \frac{C_2 - 2\varepsilon^{1/2} L}{\varepsilon^{1/2}} E_s \\ &=: \Phi(E_s) \end{aligned} \quad (63)$$

for $C_3 = \text{tr}(D) + N_x C_1^2$, $C_4 = 2L C_1^{1/2} (N_x M)^{1/4}$, and a linear function $\Phi(E_s)$. Taking expectations and using $\mathbb{E}[\Phi(E_s)] = \Phi(\mathbb{E}[E_s])$ we arrive at the integral inequality

$$\mathbb{E}[E_t] \leq \mathbb{E}[E_0] + \int_0^t \Phi(\mathbb{E}[E_s]) \, ds \quad (64)$$

and we can now apply the Gronwall lemma or comparison techniques for integral inequalities. More precisely, let $\alpha = \frac{C_2 - 2\varepsilon^{1/2} L}{\varepsilon^{1/2}} > 0$, then the time-dependent Ito's-formula implies that

$$e^{\alpha t} \mathbb{E}[E_t] \leq \mathbb{E}[E_0] + \int_0^t e^{\alpha s} \left(C_3 + \varepsilon \frac{C_4^2}{C_2} \right) \, ds \quad (65)$$

and, hence,

$$\mathbb{E}[E_t] \leq e^{-\alpha t} \mathbb{E}[E_0] + \alpha^{-1} K \quad (66)$$

with $K := C_3 + \varepsilon \frac{C_4^2}{C_2}$. Note that $\alpha^{-1} = \mathcal{O}(\varepsilon^{\frac{1}{2}})$. Hence we have shown the following

Theorem 3.4. (estimation error) *If the measurement error variance ε is chosen sufficiently small, the initial ensemble is chosen such that P_0 is invertible and the bounds of Lemmas 3.1 and 3.2 are satisfied at initial time, then the mean squared estimation error is of order $\varepsilon^{1/2}$ asymptotically in time.*

Using Markov's inequality the above estimate on the measurement error now yields for fixed t the following estimate

$$\mathbb{P}[E_t \geq c\varepsilon^q] \leq \frac{1}{c\varepsilon^q} \mathbb{E}[E_t] = \mathcal{O}(\varepsilon^{1/2-q}). \quad (67)$$

In particular, for any $q \in (0, 1/2)$ the estimation error $E_t = \|e_t\|^2/2$ is of order $\mathcal{O}(\varepsilon^q)$ with probability close to one. Note that this does not imply that for a given realization of the EnKBF, the estimation error E_t will be small all the time, i.e. that $\sup_{t \geq 0} E_t$ (or $\max_{t \in [0, T]} E_t$) is of order $\mathcal{O}(\varepsilon^q)$ with probability close to one. This latter statement requires a pathwise control, i.e. a (locally) uniform in time control of E_t , which we will derive in the next step. To this end note that (41) together with the inequality (63) imply the pathwise estimate

$$\begin{aligned} E_t &\leq e^{-\alpha t} E_0 + \frac{K}{\alpha} (1 - e^{-\alpha t}) + \int_0^t e^{-\alpha(t-s)} dM_s \\ &= e^{-\alpha t} E_0 + \frac{K}{\alpha} (1 - e^{-\alpha t}) + \alpha e^{-\alpha t} M_t + \alpha \int_0^t e^{-\alpha(t-s)} (M_t - M_s) ds, \end{aligned} \quad (68)$$

hence

$$\sup_{t \leq T} E_t \leq \left(E_0 + \frac{K}{\alpha} \right) + \sup_{t \leq T} \left(e^{-\alpha t} |M_t| + \alpha \int_0^t e^{-\alpha(t-s)} |M_t - M_s| ds \right). \quad (69)$$

In order to control the third term, first note that the quadratic variation of the martingale is given as

$$\langle M \rangle_t = \int_0^t \varepsilon^{-1} \|P_r e_r\|^2 + 2\|e_r\|^2 dr, \quad (70)$$

so that

$$\langle M \rangle_t - \langle M \rangle_s = \int_s^t \varepsilon^{-1} \|P_r e_r\|^2 + 2\|e_r\|^2 dr \leq (C_1 + 2) \int_s^t E_r dr. \quad (71)$$

In the following let $L_{T,\delta} := \sup_{0 \leq s < t \leq T} \frac{|M_t - M_s|}{(\langle M \rangle_t - \langle M \rangle_s)^{\frac{1}{2} - \delta}}$ for $\delta \in (0, \frac{1}{2})$. Theorem 5.1 in [BY82] now implies for any $\gamma \geq 1$ that there exists a finite constant $C_{\delta,\gamma}$ such that

$$\mathbb{E}[(L_{T,\delta})^\gamma]^{\frac{1}{\gamma}} \leq C_{\delta,\gamma} \mathbb{E}[\langle M \rangle_T^{\delta\gamma}]^{\frac{1}{\gamma}}. \quad (72)$$

Combining the last estimate with the previous Theorem 3.4 we obtain for $\gamma\delta \leq 1$ that

$$\mathbb{E}[(L_{T,\delta})^\gamma]^{\frac{1}{\gamma}} \leq C_{\delta,\gamma} \mathbb{E}[\langle M \rangle_T^{\delta\gamma}]^{\frac{1}{\gamma}} \leq C_{\delta,\gamma} \mathbb{E}[\langle M \rangle_T]^\delta \leq C_{\delta,\gamma} (C_1 + 2)^\delta \mathbb{E} \left[\int_0^T E_t dt \right]^\delta \leq C\varepsilon^{\frac{\delta}{2}} \quad (73)$$

for some constant C , depending on γ, δ, T, C_1 and on the bound on the mean squared error obtained in Theorem 3.4. We can therefore estimate

$$\begin{aligned} &\mathbb{E} \left[\sup_{t \leq T} e^{-\alpha t} |M_t| + \alpha \int_0^t e^{-\alpha(t-s)} |M_t - M_s| ds \right] \\ &\leq \mathbb{E} \left[\sup_{t \leq T} \left(e^{-\alpha t} \langle M \rangle_t^{\frac{1}{2} - \delta} + \alpha \int_0^t e^{-\alpha(t-s)} (\langle M \rangle_t - \langle M \rangle_s)^{\frac{1}{2} - \delta} ds \right) L_{T,\delta} \right] \\ &\leq (C_1 + 2) \mathbb{E} \left[\sup_{t \leq T} \left(e^{-\alpha t} t^{\frac{1}{2} - \delta} + \alpha \int_0^t e^{-\alpha(t-s)} (t-s)^{\frac{1}{2} - \delta} ds \right) \sup_{t \leq T} E_t^{\frac{1}{2} - \delta} L_{T,\delta} \right] \\ &\leq (C_1 + 2) \frac{\Gamma(\frac{3}{2} - \delta)}{\alpha^{\frac{1}{2} - \delta}} \mathbb{E} \left[\sup_{t \leq T} E_t^{\frac{1}{2} - \delta} L_{T,\delta} \right]. \end{aligned} \quad (74)$$

Applying Young's inequality with $p = \frac{1}{\frac{1}{2} - \delta}$ and $q = \frac{1}{\frac{1}{2} + \delta}$ we can further estimate the right hand side from above by

$$(C_1 + 2) \frac{\Gamma(\frac{3}{2} - \delta)}{\alpha^{\frac{1}{2} - \delta}} \mathbb{E} \left[\sup_{t \leq T} E_t^{\frac{1}{2} - \delta} L_{T,\delta} \right] \leq \left(\frac{1}{2} - \delta \right) \mathbb{E} \left[\sup_{t \leq T} E_t \right] + \frac{C}{\alpha^{\frac{1-2\delta}{1+2\delta}}} \mathbb{E} \left[L_{T,\delta}^{\frac{1}{\frac{1}{2} + \delta}} \right], \quad (75)$$

for some finite constant C depending on C_2 and δ . Taking expectation in (69) and using (73) to estimate the third term gives

$$\begin{aligned}\mathbb{E}\left[\sup_{t \leq T} E_t\right] &\leq \left(\mathbb{E}[E_0] + \frac{K}{\alpha}\right) + \left(\frac{1}{2} - \delta\right) \mathbb{E}\left[\sup_{t \leq T} E_t\right] + \frac{C}{\alpha^{\frac{1-2\delta}{1+2\delta}}} \mathbb{E}\left[L_{T,\delta}^{\frac{1}{\frac{1}{2}+\delta}}\right] \\ &\leq \left(\mathbb{E}[E_0] + \frac{K}{\alpha}\right) + \left(\frac{1}{2} - \delta\right) \mathbb{E}\left[\sup_{t \leq T} E_t\right] + \frac{C}{\alpha^{\frac{1-2\delta}{1+2\delta}}} \varepsilon^{\frac{\delta}{1+2\delta}}\end{aligned}\quad (76)$$

with some different constant C . Under the assumptions of Theorem 3.4, in particular $\mathbb{E}[E_0] \in \mathcal{O}(\varepsilon^{\frac{1}{2}})$, and thus $\alpha^{-1} = \mathcal{O}(\varepsilon^{\frac{1}{2}})$ for $\varepsilon \leq \varepsilon_0$, ε_0 sufficiently small, we can now find for any $\eta \in (0, \frac{1}{4})$ now a finite constant C such that

$$\mathbb{E}\left[\sup_{t \leq T} E_t\right] \leq C\varepsilon^{\frac{1}{2}-\eta}. \quad (77)$$

In particular,

$$\mathbb{P}\left[\sup_{t \leq T} E_t \geq c\varepsilon^q\right] \leq \frac{1}{c\varepsilon^q} \mathbb{E}\left[\sup_{t \leq T} E_t\right] = \mathcal{O}\left(\varepsilon^{1/2-\eta-q}\right), \quad (78)$$

which implies that for any $q \in (0, 1/2)$ the estimation error $E_t = \|e_t\|^2/2$ is of order $\mathcal{O}(\varepsilon^q)$ uniformly on $[0, T]$ with probability close to one.

4 Consistency of the ensemble Kalman-Bucy filter for linear systems

In this section, we provide a detailed analysis of the EnKBF in the case of linear model dynamics, i.e., $f(x) = Ax + b$, linear forward map, i.e. $h(x) = Hx$, full rank diffusion tensor, D , and initial ensemble, X_0^i , chosen such that P_0^M is invertible. Then the EnKBF (5) reduces to

$$dX_t^i = (AX_t^i + b)dt + D(P_t^M)^{-1}(X_t^i - \bar{x}_t^M)dt - \frac{1}{2}P_t^M H^T R^{-1} (HX_t^i dt + H\bar{x}_t^M dt - dY_t), \quad (79)$$

$i = 1, \dots, M$, from which we can extract the equation for the empirical mean, \bar{x}_t ,

$$d\bar{x}_t^M = A\bar{x}_t^M dt + bdt - P_t^M H^T R^{-1} (H\bar{x}_t^M dt - dY_t) \quad (80)$$

and the equation for the empirical covariance matrix, as defined in (6),

$$\frac{d}{dt}P_t^M = AP_t^M + P_t^M A^T + 2D - P_t^M H^T R^{-1} H P_t^M \quad (81)$$

provided P_t^M has full rank. These equations correspond exactly to the classic Kalman-Bucy filter formulas for the mean and the covariance matrix [Jaz70]. However, while one would set P_0^M and \bar{x}_0^M equal to the mean and the covariance matrix, respectively, of the given initial Gaussian distribution $N(\bar{x}_0, P_0)$ in the classic Kalman-Bucy filter formulation, the P_t^M and \bar{x}_t^M arise in our context from sampling from the initial distribution, i.e., $X_0^i \sim N(\bar{x}_0, P_0)$.

Remark 4.1. *It is well-known that solutions to (81) have full rank for all $t > 0$ even if the initial P_0^M is singular. However, note that (81) holds true only if P_0^M is non-singular and that the diffusion induced contribution in (81) needs to be replaced by $D(P_t^M)^+ P_t^M$ otherwise. This discrepancy between the Riccati equation for the classic Kalman-Bucy filter and the EnKBF is caused by our interacting particle approximation to the diffusion term in (1).*

In the following we will apply results from classical stability analysis of linear systems to our present setting. It is well-known that if (A, H) is observable, i.e., $\text{rank}[H^T, (HA)^T, \dots, (HA^{N_x-1})^T] = N_x$, and (A, C) is controllable, i.e., $\text{rank}[C, AC, \dots, A^{N_x-1}C] = N_x$, then there exists a unique positive definite solution P_∞ of the matrix Riccati equation

$$0 = AP_\infty + P_\infty A^T + 2D - P_\infty H^T R^{-1} H P_\infty, \quad (82)$$

and the solution P_t of the matrix Riccati equation

$$\frac{d}{dt}P_t = AP_t + P_tA^T + 2D - P_tH^TR^{-1}HP_t, \quad (83)$$

converges for any initial condition P_0 towards P_∞ as $t \rightarrow \infty$ with exponential rate $\lambda < \lambda_*$, where

$$\lambda_* := \inf\{-\operatorname{Re}(\lambda) \mid \lambda \text{ eigenvalue of } A - P_\infty H^T R^{-1} H\}. \quad (84)$$

(see [KS72], Theorem 4.11, and [OP96], Lemma 2.2).

The previous discussion indicates that we may expect the EnKBF to converge to the classic Kalman-Bucy filter in the case of an observable and controllable system. We will now investigate the asymptotic behavior of the EnKBF in the large ensemble size limit. We show that the empirical distribution of the EnKBF converges under appropriate conditions towards the posterior Gaussian distribution with mean and covariance determined by the Kalman-Bucy filtering equations. To this end let us first state the following a.s. result on the asymptotic behavior of P_t^M .

Proposition 4.2. *Let π_0 be the initial distribution on \mathbb{R}^{N_x} with finite second moments and invertible covariance matrix with entries*

$$\bar{P}_0(k, l) = \pi_0[x_k x_l] - \pi_0[x_k] \pi_0[x_l], \quad (85)$$

$1 \leq k, l \leq N_x$. Let X_0^i , $i = 1, 2, \dots$, be iid (π_0) , and let \bar{P}_t be the solution of the Kalman-Bucy filtering equation (83) with initial condition \bar{P}_0 . Then there exists a constant $\tilde{C} = \tilde{C}(t, A, D, H^T R^{-1} H, \max_{0 \leq s \leq t} \|\bar{P}_s\|_F, \sup_{M \geq 2} V_0^M)$ such that

$$\|P_t^M - \bar{P}_t\|_F^2 \leq e^{t\tilde{C}} \|P_0^M - \bar{P}_0\|_F^2, \quad (86)$$

where V_0^M is defined by (18) with $t = 0$.

Note that the strong law of large numbers implies that $\sup_{M \geq 2} V_0^M < \infty$ π_0 -a.s.

Proof. Using the dynamical equations (81) for P_t^M and (83) for \bar{P}_t (which of course coincides with (81)), we immediately obtain that

$$\begin{aligned} \frac{1}{2} \frac{d}{dt} \|P_t^M - \bar{P}_t\|_F^2 &\leq \langle A(P_t^M - \bar{P}_t), P_t^M - \bar{P}_t \rangle + \langle (P_t^M - \bar{P}_t) A^T, P_t^M - \bar{P}_t \rangle + 2 \langle D(P_t^M - \bar{P}_t), P_t^M - \bar{P}_t \rangle \\ &\quad - \langle P_t^M H^T R^{-1} H P_t^M - \bar{P}_t H^T R^{-1} H \bar{P}_t, P_t^M - \bar{P}_t \rangle. \end{aligned} \quad (87)$$

Using

$$\begin{aligned} &\langle P_t^M H^T R^{-1} H P_t^M - \bar{P}_t H^T R^{-1} H \bar{P}_t, P_t^M - \bar{P}_t \rangle \\ &= \langle P_t^M H^T R^{-1} H (P_t^M - \bar{P}_t), P_t^M - \bar{P}_t \rangle + \langle (P_t^M - \bar{P}_t) H^T R^{-1} H \bar{P}_t, P_t^M - \bar{P}_t \rangle \\ &\leq \|H^T R^{-1} H\|_F (\|P_t^M\|_F + \|\bar{P}_t\|_F) \|P_t^M - \bar{P}_t\|_F^2 \end{aligned} \quad (88)$$

we arrive at the following differential inequality

$$\frac{1}{2} \frac{d}{dt} \|P_t^M - \bar{P}_t\|_F^2 \leq (2\|A\|_F + \|D\|_F + \|H^T R^{-1} H\|_F (\|P_t^M\|_F + \|\bar{P}_t\|_F)) \|P_t^M - \bar{P}_t\|_F^2. \quad (89)$$

Integrating up the last inequality w.r.t. time t yields

$$\|P_t^M - \bar{P}_t\|_F^2 \leq \exp \left(2t(2\|A\|_F + \|D\|_F) + \|H^T R^{-1} H\|_F \int_0^t (\|P_s^M\|_F + \|\bar{P}_s\|_F) ds \right) \|P_0^M - \bar{P}_0\|_F^2. \quad (90)$$

In the next step we will need a uniform in M upper bound on $\|P_t^M\|_F$ that holds (locally) uniform w.r.t. time t . To this end first note that (37) implies

$$\|P_t^M\|_F \leq V_t^M \leq e^{\|A\|_F t} \left(V_0^M + \frac{\operatorname{tr}(D)}{\|A\|_F} \right), \quad (91)$$

thereby using $L_+ \leq \|A\|_F$. Since the solution \bar{P}_t of (83) is continuous, hence, also locally bounded, we can estimate the exponential in (90) from above by

$$2t \left(2\|A\|_F + 2\|D\|_F + \|R^{-1}\|_F \|H\|_F^2 \left(e^{t\|A\|_F} \left(V_0^M + \frac{\operatorname{tr}(D)}{\|A\|_F} \right) + \max_{0 \leq s \leq t} \|\bar{P}_s\|_F \right) \right)$$

which implies the assertion. \square

We can now state our main result on the asymptotic consistency of the ensemble Kalman filter.

Theorem 4.3. *Suppose that X_0^i , $i = 1, 2, 3, \dots$, are iid (π_0) where the initial distribution π_0 has finite second-order moments and invertible covariance matrix (85). Let \bar{P}_t be the solution of the Kalman-Bucy filtering equation (83) with initial condition \bar{P}_0 and \bar{x}_t be the unique solution of*

$$d\bar{x}_t = A\bar{x}_t dt + b dt - \bar{P}_t H^T R^{-1} (H\bar{x}_t dt - dY_t) \quad (92)$$

with initial condition $\bar{x}_0 := \pi_0[x]$. Then $\lim_{M \rightarrow \infty} \bar{x}_t^M = \bar{x}_t$ in L^2 , in particular in probability, for all $t \geq 0$.

Proof. Since X_0^i are iid, the strong law of large numbers implies that $\lim_{M \rightarrow \infty} P_0^M = \bar{P}_0$ π_0 -a.s. and in L^2 , since π_0 has finite second moments, thus $\lim_{M \rightarrow \infty} P_t^M = \bar{P}_t$ a.s. and in L^2 for $t \geq 0$ due to Proposition 4.2.

To see that \bar{x}_t^M converges towards the unique solution \bar{x}_t of (92) note that

$$\begin{aligned} d(\bar{x}_t^M - \bar{x}_t) &= A(\bar{x}_t^M - \bar{x}_t) dt - (P_t^M H^T R^{-1} H \bar{x}_t^M - \bar{P}_t H^T R^{-1} H \bar{x}_t) dt \\ &\quad + (P_t^M - \bar{P}_t) H^T R^{-1} dY_t \end{aligned} \quad (93)$$

and, consequently,

$$\begin{aligned} \|\bar{x}_t^M - \bar{x}_t\| &\leq \|\bar{x}_0^M - \bar{x}_0\| + \int_0^t (\|A\|_F + \|H^T R^{-1} H\|_F \|\bar{P}_t\|_F) \|\bar{x}_s^M - \bar{x}_s\| ds \\ &\quad + \int_0^t \|H^T R^{-1} H\|_F \|P_s^M - \bar{P}_s\|_F \|\bar{x}_s\| ds + \left\| \int_0^t (P_s^M - \bar{P}_s) H^T R^{-1} dY_s \right\|. \end{aligned} \quad (94)$$

Taking expectations we arrive at

$$\begin{aligned} \mathbb{E} [\|\bar{x}_t^M - \bar{x}_t\|] &\leq \mathbb{E} [\|\bar{x}_0^M - \bar{x}_0\|] + \int_0^t (\|A\|_F + \|H\|_F^2 \|R^{-1}\|_F \|\bar{P}_t\|_F) \mathbb{E} [\|\bar{x}_s^M - \bar{x}_s\|] ds \\ &\quad + \int_0^t \|H\|_F^2 \|R^{-1}\|_F \mathbb{E} [\|P_s^M - \bar{P}_s\|_F] \|\bar{x}_s\| ds + \mathbb{E} \left[\left\| \int_0^t (P_s^M - \bar{P}_s) H^T R^{-1} dY_s \right\| \right]. \end{aligned} \quad (95)$$

Using $\lim_{M \rightarrow \infty} \mathbb{E} [\|P_t^M - \bar{P}_t\|_F^2] = 0$ it follows that

$$\lim_{M \rightarrow \infty} \mathbb{E} \left[\left\| \int_0^t (P_s^M - \bar{P}_s) H^T R^{-1} dY_s \right\| \right] = 0 \quad (96)$$

by dominated convergence, and then Gronwall's lemma implies that $\lim_{M \rightarrow \infty} \mathbb{E} [\|\bar{x}_t^M - \bar{x}_t\|] = 0$. \square

5 Asymptotic limiting equations for the ensemble Kalman-Bucy equation

In this section, we will derive the non-Markovian stochastic differential equation (11) with (12) and the asymptotic dynamical equation (13) for the empirical distribution of the ensemble Kalman filter for large sample size M in the spirit of an McKean-Vlasov equation.

We first have to show now that (11) is well-posed. To this end we assume that f, h are Lipschitz continuous, that the initial condition \hat{X}_0 has finite second moments with invertible covariance matrix \mathcal{P}_0 . Recall that - given $X_t = X_t^{\text{ref}}$ - the observation process can be interpreted as Brownian motion with covariance operator R and drift term $h(X_t^{\text{ref}})$, so that we can solve (11) uniquely up to the first time τ where \mathcal{P}_τ becomes singular. Clearly, $\tau > 0$ a.s. (w.r.t. the distribution of $\{Y_s\}$). Using Itô's formula one can then show that the distribution $\hat{\pi}_t$ of \hat{X}_t indeed satisfies the McKean-Vlasov equation (13) (up to time τ).

We will show in the following couple of Lemmata that \mathcal{P}_t will stay invertible for all t , a.s. w.r.t. the distribution of $\{Y_s\}$ under appropriate assumptions on the coefficients f, h, D and R . This will in particular imply that $\tau = \infty$ a.s. which proves existence and uniqueness of a strong solution for all times t (for typical observation $\{Y_s\}$).

To this end let us first state the dynamical equations for the mean \bar{x}_t and the covariance matrix \mathcal{P}_t (analogous to (9) and (10) for the EnKBF):

$$d\bar{x}_t = \bar{f}_t dt - \mathcal{Q}_t R^{-1} (\bar{h}_t dt - dY_t), \quad t < \tau, \quad (97)$$

with $\bar{f}_t = \mathbb{E} [f(\hat{X}_t)]$ and

$$\frac{d}{dt} \mathcal{P}_t = \mathbb{E} \left[(f(\hat{X}_t) - \bar{f}_t)(\hat{X}_t - \bar{x}_t)^T + (\hat{X}_t - \bar{x}_t)(f(\hat{X}_t) - \bar{f}_t)^T \right] + 2D - \mathcal{Q}_t R^{-1} \mathcal{Q}_t^T, \quad t < \tau. \quad (98)$$

Lemma 5.1.

$$\frac{1}{N_x} \mathbb{E} \left[\|\hat{X}_t - \bar{x}_t\|^2 \right]^2 \leq \|\mathcal{P}_t\|_F^2 \leq \mathbb{E} \left[\|\hat{X}_t - \bar{x}_t\|^2 \right]^2, \quad t \leq \tau. \quad (99)$$

Proof. Similar to the proof of Lemma 2.1:

Upper bound:

$$\begin{aligned} \|\mathcal{P}_t\|_F^2 &= \sum_{k,l} \mathbb{E} \left[\left(\hat{X}_t - \bar{x}_t \right) (k) \left(\hat{X}_t - \bar{x}_t \right) (l) \right]^2 \\ &\leq \sum_{k,l} \mathbb{E} \left[\left(\hat{X}_t - \bar{x}_t \right)^2 (k) \right] \mathbb{E} \left[\left(\hat{X}_t - \bar{x}_t \right)^2 (l) \right] = \mathbb{E} \left[\|\hat{X}_t - \bar{x}_t\|^2 \right]^2. \end{aligned} \quad (100)$$

Lower bound:

$$\begin{aligned} \|\mathcal{P}_t\|_F^2 &= \sum_{k,l} \mathbb{E} \left[\left(\hat{X}_t - \bar{x}_t \right) (k) \left(\hat{X}_t - \bar{x}_t \right) (l) \right]^2 \\ &\geq \sum_k \mathbb{E} \left[\left(\hat{X}_t - \bar{x}_t \right)^2 (k) \right]^2 \geq \frac{1}{N_x} \mathbb{E} \left[\|\hat{X}_t - \bar{x}_t\|^2 \right]^2. \end{aligned} \quad (101)$$

□

Lemma 5.2. For all $t < \tau$ there exists some finite constant $C(t)$ - independent of $\{Y_s\}$ - such that

$$\sup_{0 \leq s \leq t} \mathbb{E} \left[\|\hat{X}_s - \bar{x}_s\|^2 \right] \leq C(t). \quad (102)$$

Proof. The difference $\hat{X}_t - \bar{x}_t$ satisfies the ordinary differential equation

$$\frac{d}{dt} \left(\hat{X}_t - \bar{x}_t \right) = \left(f(\hat{X}_t) - \bar{f}_t \right) + D \mathcal{P}_t^{-1} \left(\hat{X}_t - \bar{x}_t \right) - \frac{1}{2} \mathcal{Q}_t R^{-1} \left(h(\hat{X}_t) - \bar{h}_t \right) \quad (103)$$

up to time τ so that for $t < \tau$

$$\begin{aligned} \frac{d}{dt} \mathbb{E} \left[\|\hat{X}_t - \bar{x}_t\|^2 \right] &= 2 \mathbb{E} \left[\langle f(\hat{X}_t) - \bar{f}_t, \hat{X}_t - \bar{x}_t \rangle \right] + 2 \mathbb{E} \left[\langle D \mathcal{P}_t^{-1} \left(\hat{X}_t - \bar{x}_t \right), \hat{X}_t - \bar{x}_t \rangle \right] \\ &\quad - \mathbb{E} \left[\langle \mathcal{Q}_t R^{-1} \left(h(\hat{X}_t) - \bar{h}_t \right), \hat{X}_t - \bar{x}_t \rangle \right] \\ &\leq 2L_+ \mathbb{E} \left[\|\hat{X}_t - \bar{x}_t\|^2 \right] + 2 \text{tr} (D) \end{aligned} \quad (104)$$

thereby using

$$\mathbb{E} \left[\langle \mathcal{Q}_t R^{-1} \left(h(\hat{X}_t) - \bar{h}_t \right), \hat{X}_t - \bar{x}_t \rangle \right] = \|R^{-1/2} \mathcal{Q}_t^T\|_F^2 \geq 0. \quad (105)$$

This implies the same bound

$$\text{Var} (\hat{X}_t) := \mathbb{E} \left[\|\hat{X}_t - \bar{x}_t\|^2 \right] \leq e^{2L_+ t} \left(\mathbb{E} \left[\|\hat{X}_0 - \bar{x}_0\|^2 \right] + \frac{\text{tr} (D)}{L_+} \right) \quad (106)$$

as stated in Remark 2.4 for the EnKBF for $h(x) = x$ and $R = \varepsilon I$ and, therefore,

$$\sup_{0 \leq s \leq t} \text{Var} (\hat{X}_s) = \sup_{0 \leq s \leq t} \mathbb{E} \left[\|\hat{X}_s - \bar{x}_s\|^2 \right] \leq C(t) \quad (107)$$

for some finite constant $C(t)$ depending on t . Note that $C(t)$ clearly is independent of $\{Y_s\}$. □

Lemma 5.3. Let d_{\min} be the lowest eigenvalue of D . If $L^2 < 2d_{\min}\|R^{-1}\|_F\|h\|_{\text{Lip}}^2$ and if

$$\lambda_0^{\min} \geq \kappa_- := \frac{2d_{\min}\|R^{-1}\|_F\|h\|_{\text{Lip}}^2 - L^2}{2\|R^{-1}\|_F^2\|h\|_{\text{Lip}}^4 C(t)}, \quad (108)$$

then $\lambda_s^{\min} \geq \kappa_-$ for all $s < \tau \wedge t$. In particular, $\tau > t$.

Proof. We will use the representation $\lambda_t^{\min} = \inf_{\|v\|=1} \langle \mathcal{P}_t v, v \rangle$. So fix v with $\|v\| = 1$. Then

$$\frac{d}{dt} \langle \mathcal{P}_t v, v \rangle = 2\mathbb{E} \left[\langle f(\hat{X}_t) - \bar{f}_t, v \rangle \langle \hat{X}_t - \bar{x}_t, v \rangle \right] + 2\langle Dv, v \rangle - \langle R^{-1} \mathcal{Q}_t^T v, \mathcal{Q}_t^T v \rangle. \quad (109)$$

Using

$$\langle \mathcal{P}_t v, v \rangle = \mathbb{E} \left[\langle \hat{X}_t - \bar{x}_t, v \rangle^2 \right] \quad (110)$$

and

$$\begin{aligned} \langle R^{-1} \mathcal{Q}_t^T v, \mathcal{Q}_t^T v \rangle &= \langle R^{-1} \mathbb{E} \left[(h(\hat{X}_t) - \bar{h}_t) \langle \hat{X}_t - \bar{x}_t, v \rangle \right], \mathbb{E} \left[(h(\hat{X}_t) - \bar{h}_t) \langle \hat{X}_t - \bar{x}_t, v \rangle \right] \rangle \\ &\leq \|R^{-1}\|_F \mathbb{E} \left[\|h(\hat{X}_t) - \bar{h}_t\|^2 \right] \mathbb{E} \left[\langle \hat{X}_t - \bar{x}_t, v \rangle^2 \right] \\ &\leq \|R^{-1}\|_F \|h\|_{\text{Lip}}^2 \text{Var}(\hat{X}_t) \langle \mathcal{P}_t v, v \rangle, \end{aligned} \quad (111)$$

we can estimate

$$\begin{aligned} \frac{d}{dt} \langle \mathcal{P}_t v, v \rangle &\geq -2L \text{Var}(\hat{X}_t)^{\frac{1}{2}} \|v\| \langle \mathcal{P}_t v, v \rangle^{\frac{1}{2}} + 2\langle Dv, v \rangle - \|h\|_{\text{Lip}}^2 \|R^{-1}\|_F \text{Var}(\hat{X}_t) \langle \mathcal{P}_t v, v \rangle \\ &\geq -2LC(t)^{1/2} \langle \mathcal{P}_t v, v \rangle^{\frac{1}{2}} + 2\langle Dv, v \rangle - \|h\|_{\text{Lip}}^2 \|R^{-1}\|_F C(t) \langle \mathcal{P}_t v, v \rangle \\ &\geq 2d_{\min} - \frac{L^2}{\|R^{-1}\|_F \|h\|_{\text{Lip}}^2} - 2\|h\|_{\text{Lip}}^2 \|R^{-1}\|_F C(t) \langle \mathcal{P}_t v, v \rangle. \end{aligned} \quad (112)$$

Now $\lambda_0^{\min} \geq \kappa_-$ implies that $\langle \mathcal{P}_0 v, v \rangle \geq \kappa_-$ and thus $\langle \mathcal{P}_s v, v \rangle \geq \kappa_-$ for all $s < \tau \wedge t$. Hence $\lambda_s^{\min} \geq \kappa_- > 0$ for all $s < \tau \wedge t$ so that $\tau > t$, since otherwise $\lim_{s \uparrow \tau} \lambda_s^{\min} = 0$. \square

A similar analysis can be obtained for the empirical covariance matrix of the EnKBF (with constants uniformly in M) so that we have $\|(P_s^M)^{-1}\|_2 \leq C(t)$ and $\|\mathcal{P}_s^{-1}\|_2 \leq C(t)$ for all $s \leq t$ and therefore, using

$$(P_s^M)^{-1} - \mathcal{P}_s^{-1} = (P_s^M)^{-1} (\mathcal{P}_s - P_s^M) \mathcal{P}_s^{-1} \quad (113)$$

we have the estimate

$$\|(P_s^M)^{-1} - \mathcal{P}_s^{-1}\|_2 \leq C(t)^2 \|\mathcal{P}_s - P_s^M\|_2 \quad (114)$$

uniformly in $s \leq t$ for a finite constant $C(t)$ depending on t .

Lemma 5.4. Let \hat{X}_t^i , $1 \leq i \leq M$, $M \geq 2$, be the solution of (12) with initial conditions iid (π_0) and suppose that π_0 has bounded support contained in a ball with radius K . Then for all $T > 0$ there exist $\delta_0 > 0$ and $\kappa_0 > 0$ depending on T , but independent of M , such that

$$\mathbb{E} \left[e^{\delta_0 \int_0^t \frac{1}{M} \sum_{i=1}^M \|\hat{X}_s^i\|^2 ds} \right] \leq e^{2\kappa_0 \left(\frac{K^2}{M} + \|h(X_{0:T}^{\text{ref}})\|_{\infty}^2 \right)} < +\infty \quad \forall t \leq T. \quad (115)$$

Here, the expectation is taken w.r.t. the distribution of $\{Y_s\}$.

Proof. First note that Itô's formula and (12) imply that

$$\begin{aligned} d \left(\frac{1}{M} \sum_{i=1}^M \|\hat{X}_t^i\|^2 \right) &= \frac{2}{M} \sum_{i=1}^M \langle f(\hat{X}_t^i), \hat{X}_t^i \rangle dt + \frac{2}{M} \sum_{i=1}^M \langle D\mathcal{P}_t^{-1} (\hat{X}_t^i - \bar{x}_t), \hat{X}_t^i \rangle dt \\ &\quad - \frac{1}{M} \sum_{i=1}^M \langle \mathcal{Q}_t R^{-1} h(\hat{X}_t^i), \hat{X}_t^i \rangle dt - \frac{1}{M} \sum_{i=1}^M \langle \mathcal{Q}_t R^{-1} \bar{h}_t, \hat{X}_t^i \rangle dt \\ &\quad + \frac{2}{M} \sum_{i=1}^M \langle \hat{X}_t^i, \mathcal{Q}_t R^{-1} dY_t \rangle + \frac{1}{M} \text{tr}(\mathcal{Q}_t R^{-1} \mathcal{Q}_t) dt. \end{aligned} \quad (116)$$

Using Lipschitz continuity of f and h and the previous two Lemmata 5.2 and 5.3, the right hand side can be estimated from above for $t \leq T$ by

$$C(T) \left(1 + \frac{1}{M} \sum_{i=1}^M \|\hat{X}_t^i\|^2 \right) + \frac{2}{M} \sum_{i=1}^M \langle \hat{X}_t^i, \mathcal{Q}_t R^{-1} dY_t \rangle \quad (117)$$

for some uniform constant $C(T)$. Since $dY_t = h(X_t^{\text{ref}}) dt + R^{-1/2} dV_t$ we can further estimate from above for $t \leq T$

$$C(T) \left(1 + \|h(X_t^{\text{ref}})\|^2 + \frac{1}{M} \sum_{i=1}^M \|\hat{X}_t^i\|^2 \right) + \frac{2}{M} \sum_{i=1}^M \langle \hat{X}_t^i, \mathcal{Q}_t R^{-1/2} dV_t \rangle \quad (118)$$

for some possibly different constant $C(T)$. Itô's product rule now implies for $\alpha := 1 + C(T)$ and $t \leq T$

$$\begin{aligned} d \left(e^{-\alpha t} \frac{1}{M} \sum_{i=1}^M \|\hat{X}_t^i\|^2 \right) &\leq e^{-\alpha t} C(T) \left(1 + \|h(X_t^{\text{ref}})\|^2 \right) dt - e^{-\alpha t} \left(\frac{1}{M} \sum_{i=1}^M \|\hat{X}_t^i\|^2 \right) dt \\ &\quad + e^{-\alpha t} \frac{2}{M} \sum_{i=1}^M \langle \hat{X}_t^i, \mathcal{Q}_t R^{-1/2} dV_t \rangle, \end{aligned} \quad (119)$$

which implies that

$$\int_0^t e^{-\alpha s} \frac{1}{M} \sum_{i=1}^M \|\hat{X}_s^i\|^2 ds \leq \frac{1}{M} \sum_{i=1}^M \|\hat{X}_0^i\|^2 + C(T) \left(1 + \|h(X_{0:T}^{\text{ref}})\|_\infty^2 \right) + \int_0^t e^{-\alpha s} \frac{2}{M} \sum_{i=1}^M \langle \hat{X}_s^i, \mathcal{Q}_s R^{-1/2} dV_s \rangle. \quad (120)$$

To simplify notations in the following let

$$M_t := \int_0^t e^{-\alpha s} \frac{2}{M} \sum_{i=1}^M \langle \hat{X}_s^i, \mathcal{Q}_s R^{-1/2} dV_s \rangle \quad (121)$$

and observe that the quadratic variation $\langle M \rangle_t$ can be estimated from above by

$$\langle M \rangle_t = \frac{4}{M^2} \sum_{i=1}^M \int_0^t e^{-2\alpha s} \|R^{-1/2} \mathcal{Q}_s^T \hat{X}_s^i\|^2 ds \leq \frac{4 \|R^{-1/2}\|_F^2 C(T)^2}{M} \int_0^t e^{-\alpha s} \frac{1}{M} \sum_{i=1}^M \|\hat{X}_s^i\|^2 ds, \quad (122)$$

using

$$\|\mathcal{Q}_s\|_F^2 \leq \|h\|_{\text{Lip}}^2 \mathbb{E} \left[\|\hat{X}_s - \bar{x}_s\|^2 \right] \leq \|h\|_{\text{Lip}}^2 C(T)^2 \quad (123)$$

and Lemma 5.2. The assumption on the initial condition now implies for $\delta > 0$

$$\begin{aligned} \mathbb{E} \left[e^{\delta \int_0^t e^{-\alpha s} \frac{1}{M} \sum_{i=1}^M \|\hat{X}_s^i\|^2 ds} \right] &\leq e^{\delta \left(\frac{\kappa^2}{M} + C(T) \left(1 + \|h(X_{0:T}^{\text{ref}})\|_\infty^2 \right) \right)} \mathbb{E} \left[e^{\delta M_t} \right] \\ &\leq e^{\delta \left(\frac{\kappa^2}{M} + C(T) \left(1 + \|h(X_{0:T}^{\text{ref}})\|_\infty^2 \right) \right)} \mathbb{E} \left[e^{2\delta^2 \langle M \rangle_t} \right]^{1/2} \\ &\leq e^{\delta \left(\frac{\kappa^2}{M} + C(T) \left(1 + \|h(X_{0:T}^{\text{ref}})\|_\infty^2 \right) \right)} \mathbb{E} \left[e^{2\delta^4 \frac{\|R^{-1/2}\|_F^2 C(T)^2}{M} \delta \int_0^t e^{-\alpha s} \frac{1}{M} \sum_{i=1}^M \|\hat{X}_s^i\|^2 ds} \right]^{1/2}, \end{aligned} \quad (124)$$

thereby using the inequality

$$\mathbb{E} \left[e^{\delta M_t} \right] = \mathbb{E} \left[e^{\frac{1}{2} (2\delta M_t - 2\delta^2 \langle M \rangle_t)} e^{\frac{1}{2} (2\delta^2 \langle M \rangle_t)} \right] \leq \mathbb{E} \left[e^{2\delta M_t - 2\delta^2 \langle M \rangle_t} \right]^{1/2} \mathbb{E} \left[e^{2\delta^2 \langle M \rangle_t} \right]^{1/2} = \mathbb{E} \left[e^{2\delta^2 \langle M \rangle_t} \right]^{1/2}. \quad (125)$$

Hence for $\delta_0 > 0$ with

$$\delta_0 \frac{8 \|R^{-1/2}\|_F^2 C(T)^2}{M} < 1 \quad (126)$$

it follows that

$$\mathbb{E} \left[e^{\delta_0 \int_0^t e^{-\alpha s} \frac{1}{M} \sum_{i=1}^M \|\hat{X}_s^i\|^2 ds} \right] \leq e^{2\delta_0 \left(\frac{\kappa^2}{M} + C(T) \left(1 + \|h(X_{0:T}^{\text{ref}})\|_\infty^2 \right) \right)} < e^{2\kappa_0 \left(\frac{\kappa^2}{M} + \|h(X_{0:T}^{\text{ref}})\|_\infty^2 \right)} < +\infty \quad (127)$$

for a suitable $\kappa_0 > 0$. \square

From this point we can now proceed in a standard way to prove the following result:

Theorem 5.5. *Assume that $L^2 < 2d_{\min}\|R^{-1}\|_F\|h\|_{\text{Lip}}^2$. Let π_0 be a distribution on \mathbb{R}^{N_x} with finite support and invertible covariance matrix \mathcal{P}_0 satisfying $\lambda_0^{\min} \geq \kappa_-$, where κ_- is as in Lemma 5.3. Let \hat{X}_t^i be solutions of the mean-field process (11) with initial conditions $\hat{X}_0^i = X_0^i$ and X_0^i are iid (π_0) , so that the solutions \hat{X}_t^i to the mean field processes are iid too. Then*

$$\lim_{M \rightarrow \infty} \mathbb{E} \left[\frac{1}{M} \sum_{i=1}^M \|X_t^i - \hat{X}_t^i\| \right] = 0. \quad (128)$$

In particular,

$$\lim_{M \rightarrow \infty} \frac{1}{M} \sum_{i=1}^M g(X_t^i) - \hat{\pi}_t[g] = 0 \quad (129)$$

in $L^2(\mathbb{P})$, hence in probability, for any Lipschitz continuous function g .

Remark 5.6. *The last theorem implies in particular that the empirical distribution $\hat{\pi}_t^M$, defined in (8), of the EnKBF with M ensemble members converges weakly towards the distribution $\hat{\pi}_t$ of the mean field process (11) in probability w.r.t. the distribution of $\{Y_s\}$.*

Remark 5.7. *The conditions of Theorem 5.5 are satisfied for fully observed processes $h(x) = x$, measurement error covariance matrix $R = \varepsilon I$, $\varepsilon > 0$ sufficiently small, and full rank diffusion tensor D , i.e., for the filtering setting considered in Sections 2 and 3.*

Proof. Itô's formula implies that

$$\begin{aligned} d \left(\frac{1}{M} \sum_{i=1}^M \|X_t^i - \hat{X}_t^i\|^2 \right) &= \frac{2}{M} \sum_{i=1}^M \langle f(X_t^i) - f(\hat{X}_t^i), X_t^i - \hat{X}_t^i \rangle dt \\ &\quad + \frac{2}{M} \sum_{i=1}^M \langle D \left((P_t^M)^{-1} (X_t^i - \bar{x}_t^M) - \mathcal{P}_t^{-1} (\hat{X}_t^i - \bar{x}_t) \right), X_t^i - \hat{X}_t^i \rangle dt \\ &\quad - \frac{1}{M} \sum_{i=1}^M \langle Q_t^M R^{-1} (h(X_t^i) + \bar{h}_t^M) - \mathcal{Q}_t R^{-1} (h(\hat{X}_t^i) + \bar{h}_t) \rangle, X_t^i - \hat{X}_t^i \rangle dt \quad (130) \\ &\quad + \frac{2}{M} \sum_{i=1}^M \langle X_t^i - \hat{X}_t^i, (Q_t^M - \mathcal{Q}_t) R^{-1} dY_t \rangle \\ &\quad + \frac{1}{M} \text{tr} \left((Q_t^M - \mathcal{Q}_t) R^{-1} (Q_t^M - \mathcal{Q}_t)^T \right) dt \\ &= I + \dots + V. \end{aligned}$$

Using the decomposition $dY_t = h(X_t^{\text{ref}}) dt + R^{1/2} dV_t$ we can split up the stochastic integral IV into

$$\begin{aligned} \frac{2}{M} \sum_{i=1}^M \langle X_t^i - \hat{X}_t^i, (Q_t^M - \mathcal{Q}_t) R^{-1} dY_t \rangle &= \frac{2}{M} \sum_{i=1}^M \langle X_t^i - \hat{X}_t^i, (Q_t^M - \mathcal{Q}_t) R^{-1} h(X_t^{\text{ref}}) \rangle dt \\ &\quad + \frac{2}{M} \sum_{i=1}^M \langle X_t^i - \hat{X}_t^i, (Q_t^M - \mathcal{Q}_t) R^{-1/2} dV_t \rangle \quad (131) \\ &= IVa + IVb. \end{aligned}$$

We can now estimate all terms on the right hand side of the above equation for $t \leq T$, apart from the stochastic integral IVb, from above up to some process

$$U_M(t) = C_T \left(1 + \|h(X_{0:T}^{\text{ref}})\|_\infty^2 + \left(\frac{1}{M} \sum_{i=1}^M \|\hat{X}_t^i\|^2 \right)^{1/2} \right) \left(1 + \frac{1}{M} \sum_{i=1}^M \|X_t^i - \bar{x}_t^M\|^2 + \frac{1}{M} \sum_{i=1}^M \|\hat{X}_t^i - \bar{x}_t\|^2 \right), \quad (132)$$

where C_T is some finite constant, multiplied with $\frac{1}{M} \sum_{i=1}^M \|X_t^i - \hat{X}_t^i\|^2$ plus some remainder $R_M(t)$ that converges towards zero in $L^p(\mathbb{P})$ as $M \rightarrow \infty$ for all finite p . More specifically,

$$\begin{aligned} \mathbb{d} \left(\frac{1}{M} \sum_{i=1}^M \|X_t^i - \hat{X}_t^i\|^2 \right) &\leq U_M(t) \left(\frac{1}{M} \sum_{i=1}^M \|X_t^i - \hat{X}_t^i\|^2 + R_M(t) \right) dt \\ &\quad + \frac{2}{M} \sum_{i=1}^M \langle X_t^i - \hat{X}_t^i, (Q_t^M - \mathcal{Q}_t) R^{-1/2} dV_t \rangle. \end{aligned} \quad (133)$$

Indeed, this is obvious for terms I, III and V, in view of the second estimate in Lemma 5.8 below and for term II it follows from the first estimate in Lemma 5.8 below in combination with (114).

Applying Itô's product formula to the process $e^{\int_0^t U_M(s) ds} \frac{1}{M} \sum_{i=1}^M \|X_t^i - \hat{X}_t^i\|^2$ and taking expectations w.r.t. the distribution of $\{Y_s\}$, we arrive at the following estimate

$$\mathbb{E} \left[e^{-\int_0^t U_M(s) ds} \frac{1}{M} \sum_{i=1}^M \|X_t^i - \hat{X}_t^i\|^2 \right] \leq C_T \mathbb{E} \left[\int_0^t e^{-\int_0^s U_M(r) dr} U_M(s) R_M(s) ds \right] \quad (134)$$

for $t \leq T$. Since $U_M R_M$ is bounded by some finite constant plus some power of $\frac{1}{M} \sum_{i=1}^M \|\hat{X}_t^i\|^2$ and the latter one has some finite exponential moment by the previous Lemma 5.4, it follows that

$$\lim_{M \rightarrow \infty} \mathbb{E} \left[e^{-\alpha_T \int_0^t \frac{1}{M} \sum_{i=1}^M \|\hat{X}_s^i\| ds} \frac{1}{M} \sum_{i=1}^M \|X_t^i - \hat{X}_t^i\|^2 \right] = 0, \quad t \leq T, \quad (135)$$

for some $\alpha_T > 0$. Now, using Lemma 5.4 again, we also may now conclude that

$$\begin{aligned} \lim_{M \rightarrow \infty} \mathbb{E} \left[\frac{1}{M} \sum_{i=1}^M \|X_t^i - \hat{X}_t^i\| \right]^2 &\leq \sup_{M \geq 2} \mathbb{E} \left[e^{\alpha_T \int_0^t \frac{1}{M} \sum_{i=1}^M \|\hat{X}_s^i\| ds} \right] \\ &\quad \times \lim_{M \rightarrow \infty} \mathbb{E} \left[e^{-\alpha_T \int_0^t \frac{1}{M} \sum_{i=1}^M \|\hat{X}_s^i\| ds} \frac{1}{M} \sum_{i=1}^M \|X_t^i - \hat{X}_t^i\|^2 \right] = 0, \quad t \leq T. \end{aligned} \quad (136)$$

□

Lemma 5.8.

$$\|P_t^M - \mathcal{P}_t\|_F \leq 2\Sigma(t) \left(\frac{1}{M-1} \sum_i \|X_t^i - \hat{X}_t^i\|^2 \right)^{\frac{1}{2}} + R_M(t) \quad (137)$$

with $\lim_{M \rightarrow \infty} R_M(t) = 0$ a.s. and in $L^1(\mathbb{P})$. Here

$$\Sigma(t) := \left(\frac{1}{M-1} \sum_i \|X_t^i - \bar{x}_t^M\|^2 \right)^{\frac{1}{2}} + \left(\frac{1}{M-1} \sum_i \|\hat{X}_t^i - \bar{x}_t\|^2 \right)^{\frac{1}{2}}. \quad (138)$$

Similarly,

$$\|Q_t^M - \mathcal{Q}_t\|_F \leq 2(1 + \|h\|_{\text{Lip}}) \Sigma(t) \left(\frac{1}{M-1} \sum_i \|X_t^i - \hat{X}_t^i\|^2 \right)^{\frac{1}{2}} + S_M(t) \quad (139)$$

with $\lim_{M \rightarrow \infty} S_M(t) = 0$ a.s. and in $L^1(\mathbb{P})$.

Remark 5.9. Note that the factor $\Sigma(t)$ is locally bounded in t due to Lemma 5.2 and an appropriate generalization of Lemma 2.4.

Proof. First note that we can decompose

$$\begin{aligned}
P_t^M - \mathcal{P}_t &= \frac{1}{M-1} \sum_{i=1}^M (X_t^i - \bar{x}_t^M) (X_t^i - \bar{x}_t^M)^T - \mathbb{E} \left[(\hat{X}_t - \bar{x}_t) (\hat{X}_t - \bar{x}_t)^T \right] \\
&= \frac{1}{M-1} \sum_{i=1}^M \left(X_t^i - \bar{x}_t^M - (\hat{X}_t^i - \bar{x}_t) \right) (X_t^i - \bar{x}_t^M)^T \\
&\quad + \frac{1}{M-1} \sum_{i=1}^M (\hat{X}_t^i - \bar{x}_t) \left(X_t^i - \bar{x}_t^M - (\hat{X}_t^i - \bar{x}_t) \right)^T \\
&\quad + \frac{1}{M-1} \sum_{i=1}^M (\hat{X}_t^i - \bar{x}_t) (\hat{X}_t^i - \bar{x}_t)^T - \mathbb{E} \left[(\hat{X}_t - \bar{x}_t) (\hat{X}_t - \bar{x}_t)^T \right] \\
&= I + II + III.
\end{aligned} \tag{140}$$

In particular, $\|P_t^M - \mathcal{P}_t\|_F \leq \|I\|_F + \|II\|_F + \|III\|_F$. Term I can be estimated from above by

$$\begin{aligned}
\|I\|_F &\leq \left(\frac{1}{M-1} \sum_{i=1}^M \|X_t^i - \bar{x}_t^M - (\hat{X}_t^i - \bar{x}_t)\|^2 \right)^{1/2} \left(\frac{1}{M-1} \sum_{i=1}^M \|X_t^i - \bar{x}_t^M\|^2 \right)^{1/2} \\
&\leq \left(\left(\frac{1}{M-1} \sum_{i=1}^M \|X_t^i - \hat{X}_t^i\|^2 \right)^{1/2} + \sqrt{\frac{M}{M-1}} \|x_t^M - \bar{x}_t\| \right) \left(\frac{1}{M-1} \sum_{i=1}^M \|X_t^i - \bar{x}_t^M\|^2 \right)^{1/2} \\
&\leq \left(2 \left(\frac{1}{M-1} \sum_{i=1}^M \|X_t^i - \hat{X}_t^i\|^2 \right)^{1/2} + \sqrt{\frac{M}{M-1}} \left\| \frac{1}{M} \sum_{i=1}^M \hat{X}_t^i - \mathbb{E} [\hat{X}_t^i] \right\| \right) \left(\frac{1}{M-1} \sum_{i=1}^M \|X_t^i - \bar{x}_t^M\|^2 \right)^{1/2}.
\end{aligned} \tag{141}$$

Similarly,

$$\|II\|_F \leq \left(2 \left(\frac{1}{M-1} \sum_{i=1}^M \|X_t^i - \hat{X}_t^i\|^2 \right)^{1/2} + \sqrt{\frac{M}{M-1}} \left\| \frac{1}{M} \sum_{i=1}^M \hat{X}_t^i - \mathbb{E} [\hat{X}_t^i] \right\| \right) \left(\frac{1}{M-1} \sum_{i=1}^M \|\hat{X}_t^i - \bar{x}_t\|^2 \right)^{1/2}. \tag{142}$$

Finally,

$$\begin{aligned}
\|III\|_F &= \left\| \frac{1}{M} \sum_{i=1}^M \left(\hat{X}_t^i (\hat{X}_t^i)^T - \mathbb{E} [\hat{X}_t^i (\hat{X}_t^i)^T] \right) + \frac{1}{M(M-1)} (\hat{X}_t - \bar{x}_t) (\hat{X}_t - \bar{x}_t)^T \right\|_F \\
&\leq \left\| \frac{1}{M} \sum_{i=1}^M \left(\hat{X}_t^i (\hat{X}_t^i)^T - \mathbb{E} [\hat{X}_t^i (\hat{X}_t^i)^T] \right) \right\|_F + \frac{1}{M} \left\| \frac{1}{M-1} \sum_{i=1}^M (\hat{X}_t^i - \bar{x}_t) (\hat{X}_t^i - \bar{x}_t)^T \right\|_F
\end{aligned} \tag{143}$$

Adding up all terms we arrive at the estimate

$$\|P_t^M - \mathcal{P}_t\|_F \leq 2\Sigma(t) \left(\frac{1}{M-1} \sum_{i=1}^M \|X_t^i - \hat{X}_t^i\|^2 \right)^{1/2} + R_M \tag{144}$$

with the remainder

$$R_M(t) = \Sigma(t) \sqrt{\frac{M}{M-1}} \left\| \frac{1}{M} \sum_{i=1}^M \hat{X}_t^i - \mathbb{E} [\hat{X}_t^i] \right\| + \|III\|_F. \tag{145}$$

The strong law of large numbers now implies that $\lim_{M \rightarrow \infty} R_M(t) = 0$ in a.s. and in $L^1(\mathbb{P})$. The proof of the second estimate is done similarly. \square

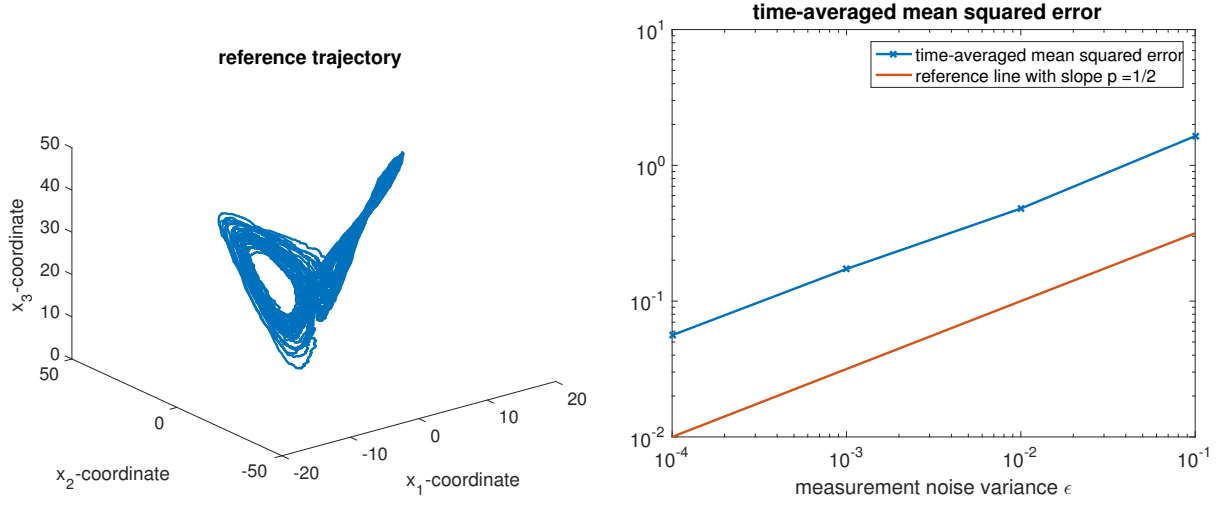


Figure 1: Reference trajectory (left panel) and time-averaged mean squared error as a function of the measurement error variance ϵ (right panel).

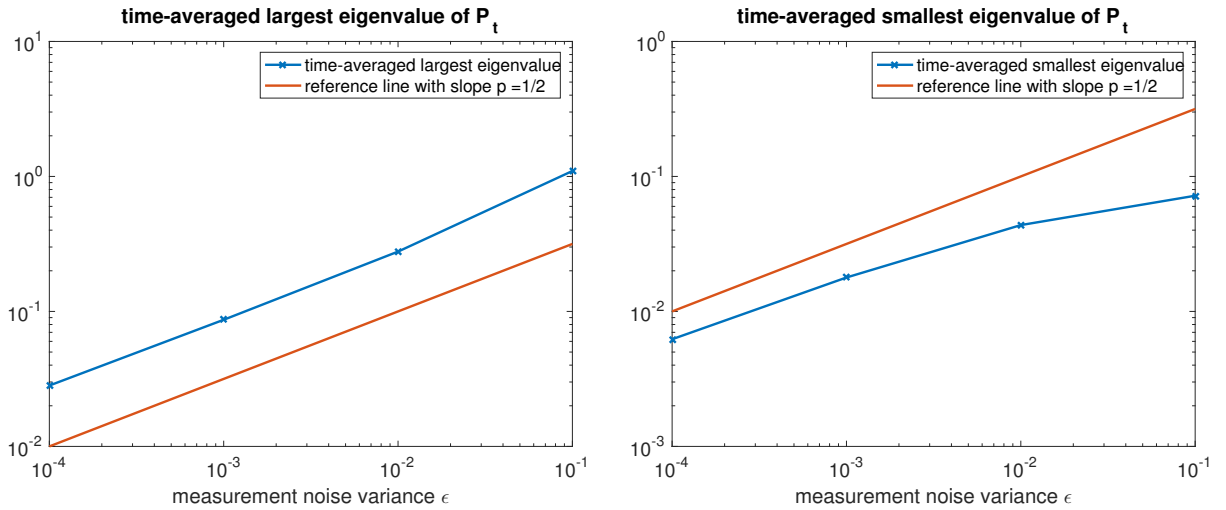


Figure 2: Time-averaged largest (left panel) and smallest (right panel) eigenvalues of P_t as a function of the measurement error variance ϵ

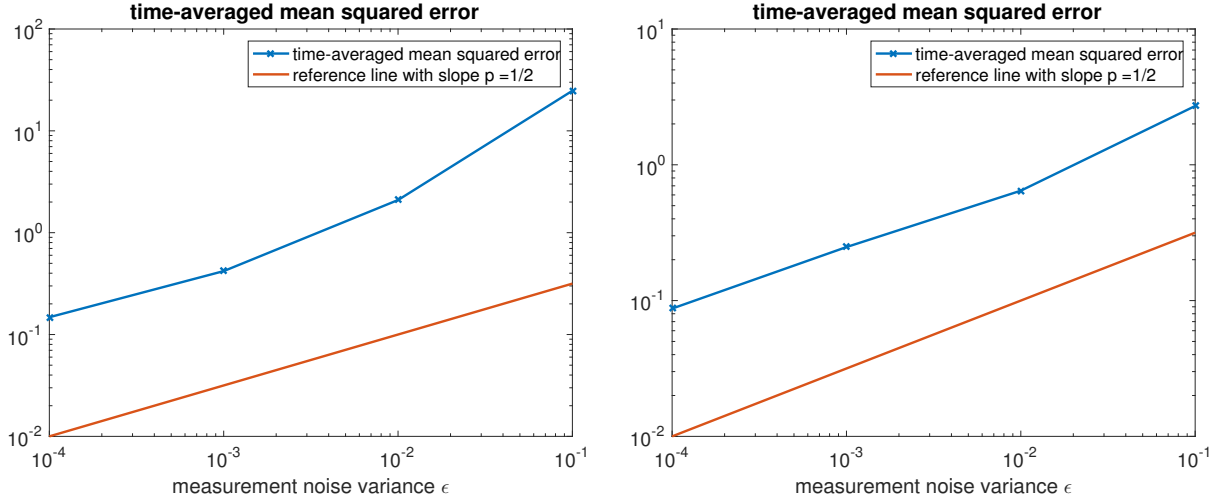


Figure 3: Time-averaged mean squared error as a function of the measurement error variance ε for ensemble sizes $M = 2$ (left panel) and $M = 3$ (right panel).

6 Numerical example

We consider the stochastically perturbed Lorenz-63 system [Lor63, LSZ15], which leads to $N_x = 3$, $D = C = I_3$, and drift term given by

$$f(x) = \begin{pmatrix} 10(x_2 - x_1) \\ (28 - x_3)x_1 - x_2 \\ x_1x_2 - \frac{8}{3}x_3 \end{pmatrix}, \quad (146)$$

where $x = (x_1, x_2, x_3)^T$. Solutions of the Lorenz-63 system diverge exponentially fast and filtering is required in order to track a reference solution. Although (146) is only locally Lipschitz continuous, the results from this paper are likely to be applicable to the Lorenz-63 system due to the existence of a Lyapunov function.

We apply the EnKBF with ensemble size $M = 4$ for values of the measurement error variances $\varepsilon \in \{0.1, 0.01, 0.001, 0.0001\}$. The stochastic evolution equations of the EnKBF are solved by the Euler-Maruyama scheme with step-size $\Delta t = 0.0005$ over a total of 10^6 time-steps. The results can be found in Figures 1 and 2. The numerical results are in agreement with our theoretical findings, which predicted an $\mathcal{O}(\varepsilon^{1/2})$ behavior of these quantities. While this scaling holds for the time-averaged mean squared error and the time-averaged largest eigenvalue of P_t for the whole range of considered values of ε , the time-averaged smallest eigenvalue truncates slightly off for the larger values of ε . We can also see that there is a gap between the smallest and largest eigenvalues of P_t on average.

We repeated the experiment for ensemble sizes of $M = 2$ and $M = 3$, in which case P_t^M is singular. We still find that the time-averaged mean squared error is roughly of $\mathcal{O}(\varepsilon^{1/2})$. See Figure 3. We will analyze the behavior of the EnKBF under singular P_t^M 's and/or partially observed phase space in a separate paper.

7 Conclusions

In this paper, we have taken first steps towards an understanding of the long-time behavior of the ensemble Kalman-Bucy filter and have derived limiting mean-field equations. Natural extensions include partially observed processes and configurations which lead to singular empirical covariance matrices P_t^M . We also plan to extend our analysis to other ensemble filter algorithms, such as the stochastically perturbed ensemble Kalman-Bucy filter and the ensemble transform particle filter. See, for example, [RC15] for more details.

References

- [AMGC02] M.S. Arulampalam, S. Maskell, N. Gordon, and T. Clapp. A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking. *IEEE Trans. Sign. Process.*, 50:174–188, 2002.

- [BC08] A. Bain and D. Crisan. *Fundamentals of stochastic filtering*, volume 60 of *Stochastic modelling and applied probability*. Springer-Verlag, New-York, 2008.
- [BR10] K. Bergemann and S. Reich. A localization technique for ensemble Kalman filters. *Q. J. R. Meteorological Soc.*, 136:701–707, 2010.
- [BR12] K. Bergemann and S. Reich. An ensemble Kalman-Bucy filter for continuous data assimilation. *Meteorolog. Zeitschrift*, 21:213–219, 2012.
- [BY82] M. Barlow and M. Yor. Aem-martingale inequalities via the garcia-rodemich-rumsey lemma, and applications to local times. *Journal of Functional Analysis*, 49:198–229, 1982.
- [DdFe01] A. Doucet, N. de Freitas, and N. Gordon (eds.). *Sequential Monte Carlo methods in practice*. Springer-Verlag, Berlin Heidelberg New York, 2001.
- [DE99] L. Dieci and T. Eirola. On smooth decompositions of matrices. *SIAM J. Matrix Anal.*, 20:800–819, 1999.
- [DMKT16] P. Del Moral, A. Kurtzmann, and J. Tugaut. On the stability and the uniform propagation of chaos of extended ensemble Kalman-Bucy filters. Technical Report arXiv:1606.08256v1, INRIA Bordeaux Research Center, 2016.
- [Eve06] G. Evensen. *Data assimilation. The ensemble Kalman filter*. Springer-Verlag, New York, 2006.
- [GMT11] F. Le Gland, V. Monbet, and V.D. Tran. Large sample asymptotics for the ensemble Kalman filter. In *The Oxford Handbook of Nonlinear Filtering*, pages 598–631, Oxford, 2011. Oxford University Press.
- [GTH13] C. González-Tokman and B.R. Hunt. Ensemble data assimilation for hyperbolic systems. *Physica D*, 243:128–142, 2013.
- [Jaz70] A.H. Jazwinski. *Stochastic processes and filtering theory*. Academic Press, New York, 1970.
- [KLS14] D. T. Kelly, K. J. H. Law, and A. Stuart. Well-posedness and accuracy of the ensemble Kalman filter in discrete and continuous time. *Nonlinearity*, 27:2579–2604, 2014.
- [KM15] E. Kwiatowski and J. Mandel. Convergence of the square root ensemble Kalman filter in the large ensemble limit. *SIAM/ASA J. Uncertainty Quantification*, 3:1–17, 2015.
- [KS72] Kwakernaak and Sivan. *Linear Optimal Control Systems*. Wiley Interscience, New York, 1972.
- [Lor63] E.N. Lorenz. Deterministic non-periodic flows. *J. Atmos. Sci.*, 20:130–141, 1963.
- [LSZ15] K. Law, A. Stuart, and K. Zygalakis. *Data Assimilation: A Mathematical Introduction*. Springer-Verlag, New York, 2015.
- [MH12] A. Majda and J. Harlim. *Filtering Complex Turbulent Systems*. Cambridge University Press, Cambridge, 2012.
- [Mor13] P. Del Moral. *Mean field simulation for Monte Carlo integration*. Chapman & Hall, London, 2013.
- [Øks00] B. Øksendal. *Stochastic Differential Equations*. Springer-Verlag, Berlin-Heidelberg, 5th edition, 2000.
- [OP96] D. Ocone and E. Pardoux. Asymptotic stability of the optimal filter with respect to its initial condition. *SIAM J. Control and Optimization*, 34:226–243, 1996.
- [RC15] S. Reich and C.J. Cotter. *Probabilistic Forecasting and Bayesian Data Assimilation*. Cambridge University Press, Cambridge, 2015.
- [Rei11] S. Reich. A dynamical systems framework for intermittent data assimilation. *BIT Numer Math*, 51:235–249, 2011.
- [Rel69] F. Rellich. *Perturbation Theory for Linear Operators*. Gordon and Breach, New York, 1969.
- [TMK16] X.T. Tong, A.J. Majda, and D. Kelly. Nonlinear stability and ergodicity of ensemble based Kalman filters. *Nonlinearity*, 29(2):657, 2016.